# AUTOMATED COMPENSATION REPORTS USING ONLINE SURVEYS AND COLLABORATIVE FILTERING

## CROSS REFERENCE TO RELATED APPLICATIONS

[0001]    This application is an application filed under 35 U.S.C. § 111(a), claiming benefit pursuant to 35 U.S.C. § 120 of the filing date of the Provisional Application Serial No. 60/436,809 filed on December 27, 2002, pursuant to 35 U.S.C. § 111(b), which is incorporated herein, in its entirety, by this reference thereto.

## BACKGROUND OF THE INVENTION

### Field of the Invention

[0002]    The invention relates to correlating statistical records. More particularly, the invention relates to correlating compensation records to unique individual profiles.

### Background of the Invention

[0003]    Today, many reports are available that allow a user to find, read, purchase, or otherwise acquire reports on worker compensation. Most often these reports indicate average pay rates by industry, job type, locale, and sometimes they report more specific information about a particular industry or job, such as bonuses, stock options, average workweek, or immigration status, among other things. To create such compensation reports two approaches are typically used. One such approach is for a human analyst to research and find a statistically valid number of individuals with like characteristics, and devise a suite of compensation reports. This process is tedious, labor intensive, and often expensive. For a truly detailed report the analyst must be relied on to do substantial investigation and synthesize and apply this information to the case at hand. Compensation consultants with years of experience and resources can generally accurately profile an individual's worth in the market place, however such an analysis is extremely specialized and out of the reach of the typical consumer. Simpler and less costly reports are available but they are generally broadly classed and offer little utility.

[0004]    The majority of software-based analysis provides a less expensive alternative but yields correspondingly limited information. Compensation services using

current computer analysis programs generally gather data using some form of questionnaire and then feed the appropriate data into a computer database or spreadsheet. More typically, generalized data, such as from the US Bureau of Labor and Statistics, is used as a base and then extrapolated based on region and date, and often combined with third party surveys. Typically, a computer then is instructed to run an analysis of the data to provide statistical information, such as averages, medians, and standard deviations on pre-determined groups of people. However the information provided is not unique to an individual, but instead, is a conglomeration of data that the program determines best represents the individual. Because the categorization of the individual is based solely on a limited, predetermined set of responses to the questionnaire, it offers little to no opportunity for evaluating unique characteristics. For example, an automated compensation service may categorize and calculate data showing that the average yearly salary of a "Computer Programmer Level 3" in Washington state is $64,250. This may or may not be applicable to a "Senior Application Software Engineer" with "10 years of experience" and special training in the skill "C++," but because the closest answer describing the Senior Application Software Engineer's position in the initial survey was a "Computer Programmer Level 3," the Senior Application Software Engineer has thus been categorized ineffectively, which removes any unique abilities that he may possess.

[0005]    The Senior Application Software Engineer reading the aforementioned report cannot be sure how closely the published report figures apply to himself individually. There are a multitude of factors that affect any one individual's job compensation. The current generalized reporting methods for compensation reports, cannot and do not incorporate factors that provide for an accurate job comparison and compensation analysis for individual users. Today's methods require the user to gauge or self-approximate themselves to a group of people being reported. Typically, such approximations are grouped by a specific job title that a human compensation analyst predetermines when creating a report, or when designing a computer service that eventually generates the report. This grouping is generally not an exact match with the user's actual job title and responsibilities and often has little applicability to the users individual qualities. For example, the compensation analyst might have created a report for an isolated group called Computer Programmer Level 3. For individuals who possess the same characteristics as the data files used to create this group, the

reports generated from such a compensation analysis are reasonably accurate. However, for individuals possessing unique capabilities, experiences, skills, or talents, the reports are essentially useless. The data are by definition misapplied because any differences in the compared data are arbitrarily reflected in the compensation report. This introduces doubt on the user's part as to how closely he can trust the report's applicability.

[0006]    Current compensation analysis techniques do not provide users with affordable, accurate, and personalized compensation reports. Job specific variables, critical to the accurate assessment of an individual's worth, are not correctly identified or uniformly applied.  Furthermore individuals within a particular field are unaware of the value of certain, often easily obtainable, qualifications. There is a need, therefore, for an apparatus and method that provides online compensation reports using a more flexible survey system to produce dynamic profiles based on unique individual attributes, and to provide automated comparisons and reports that account for these attributes.

## SUMMARY OF THE INVENTION

[0007]    In one embodiment of the invention, profiles are used to produce individualized compensation reports. A survey engine is used to produce profiles of individuals that identify the individuals' unique characteristics. The survey engine incorporates a collaborative filtering engine that determines appropriate questions to ask the user during the survey, and also provides suggested possible answers. Additionally, the system allows for the use of open-text questions that allow for new answers to be input by the user, without the prior need for an administrator to pre-define the possible values for the system, as is typical in prior art. The system incorporates affinity groups around profile attributes (question answers), providing a basis for gauging similarity of profiles for various comparison and aggregation purposes. A collaborative filtering engine is incorporated, which is both periodically modified by an administrator, and also tuned by users themselves based on their actions and responses. New affinity groups (associations of profiles) are incorporated by the survey engine to suggest new questions and possible answers in a survey. Additionally, some affinity groups are generated automatically by the system, and finally by users themselves to create new, interesting relationships among profiles.

[0008] The collaborative filtering engine enables the capture of profile attributes that are targeted compensation variables concerning a profile. The system automatically incorporates new profiles into existing affinity groups, which allows more targeted survey questions and possible answers to be determined, without requiring constant human training or intervention.

[0009] Without a collaborative filtering system, the system would be unable to administer surveys accurately for users who do not fit into pre-defined categories because it is currently impossible to categorize every occupational variation. Because the system allows for open-text answers to questions, and because the system provides for different questions to be automatically determined and asked for differing types of user job profile, the collaborative filtering system, along with affinity groups and other requirements, is employed. Because the collaborative filtering system allows for the system to make educated guesses within defined constraints, the system can handle new categorizations more effectively than a system wholly defined by a human administrator. Additionally, in this system, the administrator defines constraints that prohibit the survey from asking obviously wrong or out of place questions. An example of a constraint is requiring that if the user does not answer any of the suggested questions, a default question is always asked.

[0010] One aspect of using a collaborative filtering system to define a survey is that the system accommodates a much larger population of data using a more targeted survey. Previous implementations of compensation surveys relied on a smaller sample size, a broad survey with generalized questions, or a larger base of analysts to design and conduct surveys and categorize the data. Thus, the survey was constrained purely by the human resources needed to conduct it. The invention described herein is not constrained as such and requires far fewer human resources to conduct such a detailed survey across many different job categories.

[0011] Another aspect of the invention comprises the ability to search through the data that has been collected by the survey engine, using detailed search criteria defined by a search definition document. The document applies a scoring and filtering mechanism that returns the most appropriate set of profiles based on the user's goal for the analysis. The search document is useful because it allows an administrator to define natural relationships between profile attributes, such as those that apply routinely to the realm of compensation analysis of various classes of profiles. The

document is then interpreted by a software algorithm, and used during the retrieval of relevant profiles, which can then be used for tallying and reporting. Multiple search definition documents can be created quickly, each one for a different goal. For example, one search goal may be to weight skills and certifications very highly. The results are useful for analyzing how people with similar technical skills are compared to each other. Another search definition document could weight experience and education higher, which is useful in seeing how a user compares to those profiles who have similar experience and education levels. A third example could be to hold location constant, and compare a user to all other profiles with matching attributes in the same location.

[0012]    Another aspect of the invention comprises the automated ability to summarize and present the results of a profile search in a format that is useful to a human for compensation comparisons. A Chart is defined as a series of values, such as skills paired, with a series of measures, such as average salary, median salary, standard deviation, etc. A sample chart could be called "Average Salary By Skill" and it would list each Skill along with an associated average salary. A Report is defined as a series of charts combined to provide an overall picture and analysis for a user. For instance, consider a report that aims to discover how a user compares with regard to skills and experience in similar jobs. This report incorporates many charts which are combined into a format that gives a user a good analysis and understanding of the results of the user's search goal.

[0013]    The prior art approaches are unable to automate the selection of charts within a compensation report, such as determining if "Average Salary by Practice Area" is applicable to a report presented for a particular user profile. It may only be applicable to a Lawyer, for instance.  The prior art typically requires that the administrator know in advance, every chart that would be relevant for a particular user. This restricts known approaches in one of two ways: either they limit the number of charts available, such that all charts are available to all compensation reports, regardless of their relevance, or they predefine different compensation reports by industry or job category, which would require a large amount of labor. In contrast, the invention described herein determines if a chart is relevant to a user based on the user's own profile, and displays it, and also only shows it if enough data exists for it to be valid statistically. An advantage of the invention herein is provided because there

may be thousands of different charts of many different types, based on many different attributes and measures, and only a subset of them may apply to a particular users' compensation analysis. For instance, a Lawyer may wish to see a report concerning "Average Bonus by Number of Hours Billed per Year," but this report may not be applicable to Teachers or CEOs. In the prior art, each of these compensation reports must be compiled by a human analyst, created at expense, or details are ignored, leaving only the most common charts in the compensation report, such as "Average Salary by Occupation and Location," which are often the least useful to an individual trying to compare his compensation against his peers. The invention herein allows for a more targeted and relevant compensation report by focusing in an automated, scalable fashion, on attributes that are unique to a particular user, in addition to the attributes that are most common to all users.

[0014]    The invention may also be used to match individual profiles to resumes because the profiles described herein are typically a subset of data gathered for a resume. For example, a user entering their profile into the system described herein could be easily matched to resumes using the same mechanisms employed to match against other profiles.

## BRIEF DESCRIPTION OF THE DRAWINGS

[0015]    Figure 1 is an overview of a website implementation of a system according to the invention;

[0016]    Figure 2 is a general user flow through the system according to the invention;

[0017]    Figure 3 is a specific user flow for the first time a user accesses the system according to the invention;

[0018]    Figure 4 is a flow for the system to suggest a FieldGroup to a user according to the invention;

[0019]    Figure 5 is a flow for saving a user's answer to a database according to the invention;

[0020]    Figure 6 is an exemplary profile according to the invention;

[0021]    Figure 7 is an example of the system suggesting a FieldGroup in a survey according to the invention;

**[0022]** Figure 8 is a continuation of the example shown in Figure 7 according to the invention;

**[0023]** Figure 9 is an example of the system suggesting popular answers for a FieldGroup, in the survey according to the invention;

**[0024]** Figure 10 is a continuation of the example shown in Figure 9 according to the invention;

**[0025]** Figure 11 is a flow for calculating and summarizing data according to the invention;

**[0026]** Figure 12 is a flow for storing tokenized data in a database according to the invention;

**[0027]** Figure 13 is a flow for populating an affinity group according to the invention;

**[0028]** Figure 14 is a flow for generating a custom report for a user according to the invention;

**[0029]** Figure 15 is an overview of the different aspects of a profile search according to the invention;

**[0030]** Figure 16 is a flow for a profile search according to the invention;

**[0031]** Figure 17 is an example of an affinity group according to the invention;

**[0032]** Figure 18 is a flow for the rules engine according to the invention;

**[0033]** Figure 19 is a series of rules used in the rules engine according to the invention;

**[0034]** Figure 20 is a screen shot of the Home Page according to the invention;

**[0035]** Figure 21 is a screen shot of the Suggest Question User Interface according to the invention;

**[0036]** Figure 22 is a screen shot of the Browse AnswerValue User Interface according to the invention;

**[0037]** Figure 23 is a screen shot of the Confirm Answer User Interface according to the invention;

**[0038]** Figure 24 is a screen shot of the Choose Compensation Questions User Interface according to the invention;

**[0039]** Figure 25 is a screen shot of the Salary FieldGroup User Interface according to the invention;

**[0040]** Figure 26 is a screen shot of the Run Report User Interface according to the invention;

**[0041]** Figure 27 is a screen shot of the Summary Report User Interface according to the invention;

**[0042]** Figure 28 is a screen shot of the Profile User Interface according to the invention;

**[0043]** Figure 29 is a screen shot of the Smart Report Page 1 User Interface according to the invention;

**[0044]** Figure 31 is a screen shot of the Smart Report Page 2 User Interface according to the invention;

**[0045]** Figure 32 is a screen shot of the Smart Report Page 3 User Interface according to the invention;

**[0046]** Figure 33 is a screen shot of the Smart Report Page 4 User Interface according to the invention;

**[0047]** Figure 34 is a screen shot of the Smart Report Page 5 User Interface according to the invention;

**[0048]** Figure 35 is a screen shot of the Smart Report Page 6 User Interface according to the invention;

**[0049]** Figure 36 is a screen shot of the Research Center User Interface according to the invention;

**[0050]** Figure 37 is a screen shot of the Alerts User Interface according to the invention;

**[0051]** Figure 38 is an XML specification for FieldGroups according to the invention;

**[0052]** Figure 39 is an XML specification for Fields according to the invention;

**[0053]** Figure 40 is an XML specification for Affinity Definition according to the invention;

**[0054]** Figure 41 is an XML specification for Chart according to the invention;

**[0055]** Figure 42 is an XML specification for Profile according to the invention;

**[0056]** Figure 43 is an XML specification for AnswerGroup according to the invention;

**[0057]** Figure 44 is an XML specification for ReportGroup according to the invention;

**[0058]** Figure 45 is an XML specification for ReportArea according to the invention;

**[0059]** Figure 46 is an XML specification for Member Report according to the invention;

**[0060]**   Figure 47 is an XML specification for Wizard according to the invention;

**[0061]**   Figure 48 is an XML specification for Relation according to the invention;

**[0062]**   Figure 49 is an XML specification for Level according to the invention;

**[0063]**   Figure 50 is an XML specification for Profile Search according to the invention;

**[0064]**   Figure 51 is an XML specification for Matchgroup according to the invention;

**[0065]**   Figure 52 is an XML specification for Ranking according to the invention;

**[0066]**   Figure 53 is an XML example for FieldGroup/Field according to the invention;

**[0067]**   Figure 54 is an XML example for Aggregate Definitions according to the invention;

**[0068]**   Figures 55A-55C are an XML example for ReportGroup/ReportArea according to the invention;

**[0069]**   Figures 56A-56C are an XML example for Surveys according to the invention;

**[0070]**   Figure 57 is an XML example for Answer Relations according to the invention;

**[0071]**   Figure 58 is an XML example for FieldGroup Relations according to the invention;

**[0072]**   Figure 59 is an XML example for Levels according to the invention; and

**[0073]**   Figures 60A-60B are an XML example for Profile Search.

## DETAILED DESCRIPTION OF THE INVENTION

**[0074]**   A method and apparatus for providing targeted online compensation reports that accounts for unique individual job characteristics by using dynamic profiles is described in detail herein. In the following description, numerous specific details are provided for survey flow, affinities, levels, suggest FieldGroup, suggest popular answers, save answers, profile search, scoring system, report aggregation, and rules engine. One skilled in the relevant art, however, will recognize that the invention can be practiced without one or more of the specific details, or with other symbols, methods, etc. In other instances, well-known methods or techniques are not shown, or are not described in detail, to avoid obscuring aspects of the invention.

## Definitions

[0075]     Following are definitions used throughout the document:

[0076]     Field – A single piece of information, corresponding to a particular question asked by the system, and further as shown in Fig. 39.  Examples include, "City," "State," "Bonus Amount," "Skill," and "Job Title.

[0077]     FieldGroup – A set of related Fields that form a logical grouping of information into a single record, and further as shown in Fig. 38. Also often referred to as a Question. Examples include "Salary"- a FieldGroup consisting of "Salary Amount," "Currency" and "Average Workweek."  "Job Location" is a FieldGroup consisting of "City," "State," and "Country."

[0078]     Survey – A set of FieldGroups asked in a particular order, and as further shown in Fig. 56. An administrator may fix the order, or the system may determine the order using the "Suggest FieldGroup" algorithm defined below.

[0079]     AnswerValue (also Value, Answer or Attribute) – A value (piece of information) corresponding to a Field, used as a response.  Examples include for the for field "City," AnswerValues could be "Seattle," "San Francisco," "Miami," and "Paris."

[0080]     Profile – A set of FieldGroups, Fields, and AnswerValues that form a logical representation of an individual person or group of people, and further as shown in Fig. 42.  Examples include  "Industry=Finance, Job=Accountant, Employer-Name=Ernst & Young, Certification=CPA, Certification=CMA, Specialty=Taxation, Specialty=Cost Accounting, Years-In-Field=12, Salary-Amount=60,000, Salary-Currency=US Dollars, Salary-Workweek=40 Hours, Location-City=Baltimore, Location-State=Maryland, Location-Country=USA, School-Name=Princton, School-Degree=Masters, School-DegreeYear=1995, Age=43, etc."

[0081]     Affinity Group (or Affinity) – A grouping of profiles that share common profile attributes, and further as shown in Fig. 17.  An example is "All People who work in Law or Legal Professions."

[0082]     Affinity Definition – A Boolean representation of the common attributes shared by an affinity group.  Examples for  "All People who work in Law or Legal Professions," the exemplary affinity definition include "Job=Lawyer OR Job=Attorney OR Job=District Court Judge or Industry=Legal OR Specialty=Trial Law OR etc."

[0083]     Profile Search – A detailed set of criteria used for matching profiles to other profiles and creating a scoring system that ranks the validity of the match.

**[0084]** Profile Search Document – A document that encompasses the criteria defined for the profile search.

**[0085]** Chart – An aggregation of data limited to an affinity group, in a format understandable to an end user. An example is a Bar Chart for "Average Salary By Specialty."

**[0086]** Report – A set of charts combined in a specific layout to provide a detailed analysis of a profile search and comparison goal, as shown in Figs. 20 through 36.

## Overview

**[0087]** Every individual possesses unique distinguishing characteristics in their employment profile. These unique characteristics, even when seemingly minor, can correspond to differences in employment compensation. Being able to have a custom report that compares relevant characteristics for each individual, and having an understanding of what the market is willing to pay for the individual's abilities based on such comparison, is an important step in finding and effectively negotiating employment opportunities as well as making informed career decisions. The method and apparatus presented herein provide comparative compensation reports based on characteristics determined through a survey and a scored-attribute-matching search and reporting process.

**[0088]** Reference is now made to Fig. 2 where a general user flow through the system is shown. A user begins the investigation of their worth (200) by accessing an Internet Web site through a user interface, such as a personal computer, personal data assistant, or similar device (Fig. 1). Once access has been established, the user conveys to the system desired compensation comparison objectives/goals (202), as shown in Fig. 2. Once the objectives have been identified a survey engine (also referred to as survey wizard, see Fig. 3) 12 (Fig. 1) begins identifying the unique characteristics of the user that are most applicable to determining a compensation level, as shown in Fig. 21. The survey engine uses a combination of open and closed FieldGroups to create an individual profile. An open FieldGroup is defined as having at least one Field for which the user may enter an answer in free text. A closed FieldGroup is defined as having a Fields for which the user may only select from a list of predefined choices. After each FieldGroup is answered during the survey process, the individual profile is associated with one or more affinity groups. The affinity groups

are discrete groupings of individual profiles based on Boolean criteria. The Boolean criteria, either simple or complex, are based on a combination of values from one or more attributes on a profile, such as job, skill, location, etc. A user's profile attributes are compared against each affinity definition and, if the profile attribute meets the affinity definition criteria, the user is considered to be a member of that affinity group. For example, a user may be considered part of the affinity group "People in Information Technology/Computer Networking" if they have answered having certain skills, such as Skill=TCP/IP or Skill=Cisco Routers or Skill=Windows NT Networking or Certification=Microsoft Certified Systems Engineer.

[0089] The survey engine, through iteration, asks FieldGroups shown in Fig. 21 of the user until the engine determines no more FieldGroups should be asked. Upon beginning the survey process, the user's objectives/goals are confirmed through an initial set of questions. The goal establishes broad areas that should be investigated by the survey engine. The engine, using this goal, suggests the next FieldGroup (Question) to be posed to the user, from the set of all available FieldGroups in the system shown in Fig. 4. The FieldGroup along, with possible popular answers obtained from previous questioning of different users, are presented to the user as shown in Figs. 21 through 23. The system examines the answer offered by the user and, if appropriate, saves it in a database as shown in Fig. 5. This process continues until the engine determines that the desired goal has been ascertained and no more FieldGroups need to be asked. The engine queries if any more goals need to be examined returning the process to establishing a goal for Questioning. The process continues as described herein until profile data for all the applicable goals are created and the survey is complete.

[0090] Reference is now made to Fig. 40 where a specification for an affinity definition is shown. The system uses a concept called "affinities" or "affinity groups" to categorize and group users in many segments of the application. An affinity group is defined as a group of profiles defined by a set of profile criteria, called an affinity definition. Affinity definitions are a combination of values from fields defined in the system also shown in Fig. 6. For example, Job="Human Resource Manager" is an affinity definition for the affinity group consisting of all the profiles such that users answered Job="Human Resources Manager." Criteria can be logically defined to create more complex criteria, using standard Boolean operators, such as AND and

OR. For example, the affinity group called "San Francisco Java Programmers" might have the following affinity definition: "City=San Francisco AND (Job=Software Programmer or Job=Software Developer or Job=Web Developer) AND Skill=Java." The corresponding set of profiles that match this definition is the affinity group. The affinity definition is stored in a relational format, which is easily and quickly retrievable, and searchable. One skilled in the art would recognize that such a definition could be stored in a wide variety of formats suitable to the task at hand, or modified for improved CPU performance. A program quickly compares the user's profile and compares it to all affinity definitions stored in the system to determine to which affinity groups a user belongs. One skilled in the art would be able to recognize that more complicated affinities are possible, using more complex Boolean logic, such as NOT, XOR, etc. Additionally, affinities can be created which group numeric data by ranges. For instance, all people in a certain age range could be grouped together. Affinities are associations of data, defined by a human, an administrator, or users, which allow the system to create a more intelligent output for use by many of the sub-systems described herein.

[0091] The survey engine is required to ask FieldGroups that are relevant for a particular user. Many different job profiles have different attributes that affect one's compensation. For example, a CEO may need an answer to the FieldGroup (question) Company Revenue in his/her profile, while a Lawyer might need Practice Area or Bar Association Memberships. The suggest FieldGroup algorithm is designed to ask pertinent questions of the user depending on the user's compensation analysis goal and the type of profile the user has, learned through iterative questioning. Open-text questions provide a unique challenge in automated systems because an administrator cannot anticipate them. For instance, prior art would allow for a selection of an industry from a drop-down list box containing a set of known values. Because all industries would be known to the system, logic could be added to recommend a new question based upon the answer to the previous question. In such a system, a complex decision tree is created in advance, and the system is able to create a custom survey based upon the answers given by the user. However, these systems are subject to the limitation of knowing all possible answers, and spending a significant amount of labor to program the decision tree to account for all possibilities. It also makes it very difficult to add new questions because the added responsibility of mapping each

answer to the decision tree must be part of the process. The invention described herein is not subject to such limitations because it allows the entry by users of any answer, as further shown in Fig. 23. Typically, this open-ended question is enabled by a text box with few restrictions on what can be entered, versus a drop down box, where a user must select one of the pre-defined answers.

[0092]   Because the system allows open-ended text answers, a standard expert system is not possible, as it is impossible to create a decision tree for an infinite number of questions that would make up a survey. In such situations in other art, an artificial intelligence mechanism is usually employed to deal with the unknown variances in user input concerning which questions the users answer, as well as what those answers are.  The invention requires that a high degree of recall is maintained, which means a new user coming and using the system with the same set of profile data as an existing user should have the same set of FieldGroups recommended in the same order (deterministic). To accommodate these requirements, *i.e.,* both recall and variability, the system defines relationships between FieldGroups, as further shown in Fig. 58, a minimum set of required FieldGroups, as further shown in Fig. 55, and an order that the FieldGroups appears during the survey, as further shown in Fig. 59. These system rules are combined with a set of weights and a subsystem which allows a qualified guess as to which FieldGroups should be suggested next, without requiring a predefined survey for each unique type of job profile.

[0093]   To pick the next FieldGroup to be asked of the user, the survey engine selects the most popular FieldGroup that the user has not yet answered but that is related to the FieldGroups already posed to the user, and which is also contained in the profiles of the affinity group(s) to which the user currently belongs, subject to the level constraints and FieldGroup relationships described herein. An administrator establishes the relationships and levels for FieldGroups previous to a user completing the survey. As explained earlier, the system uses a collaborative filtering architecture for the survey engine, to allow for a large number of FieldGroups, and minimal administrative input. Although a collaborative filtering architecture is used, an alternative architecture, such as a neural network can also be employed. Administrative tasks related to Suggesting FieldGroups are limited to cleaning input data that are used to teach the system, defining FieldGroup Relations, as shown in Fig. 58, defining FieldGroup Levels, as shown in Fig. 59, and defining affinity groups.

**[0094]** The inputs to the survey engine are a user's previous answers as shown in Fig. 28, affinity groups as shown in Fig. 17, relationships between FieldGroups (FieldGroup relations document) as shown in Fig. 58, FieldGroup levels (or FieldGroup priority), and a triplet called "Popular FieldGroups," that consists of an affinity group, a FieldGroup, and a weighted value. The output for the survey engine is a single suggested FieldGroup. The constraints in the suggested FieldGroup system consist of FieldGroup levels or priorities as shown in Fig. 59 and defaults. The survey wizard prioritizes the related FieldGroups by first selecting FieldGroups that possess the highest level (or priority) as defined by a level document for the particular wizard or broad goal. Each FieldGroup is assigned one and only one level per wizard. Each level is assigned as a positive integer value to a FieldGroup, with lower levels constraining FieldGroups to appear earlier in the survey, and higher levels constraining the FieldGroup to appear later in the survey. This ensures that certain FieldGroups appear at or near the beginning of the survey and other FieldGroups appear at or near the end. Additionally, the level document also groups FieldGroups together in each level, and assigns one FieldGroup from the group as a default FieldGroup, so that if an administrator wanted to ensure that at least one FieldGroup out of a group of FieldGroups is always asked, it is asked if no other FieldGroup from that level is asked. For instance, there may be two separate FieldGroups, "Job" and "Position." Position may be a specialized type of job, which would be appropriate to ask in certain industries, such as for a professional baseball player. In that case, "job" need not be asked because that information is captured in the "position" Fieldgroup. For the most part, most profiles only need to answer "job," whereas in specialized cases they may need to answer alternate FieldGroups such as "position." Therefore "job" and "position" are grouped, and "job" is set by an administrator as the "default" FieldGroup because the system requires that if people do not answer any other job-like question they must answer job.

**[0095]** The weighted values on the Popular FieldGroups triplet are stored in a relational format. The weights represent each FieldGroup's popularity. This is defined as the number of profiles for each affinity group that have answers for that FieldGroup. Popularity is calculated by tallying up, for each FieldGroup, the number of users who have answered that FieldGroup for each affinity group, and an association between an affinity, a FieldGroup, and a popularity, and is stored in a relational table. By this

mechanism, a feedback loop is created between the user's profile questions and answers, and the survey wizard. As user profiles are entered into the system via the survey process, new associations of similar profiles are produced and those results are integrated into the popularity of each corresponding Fieldgroup, yielding a subsequently more and more precise survey for differing types of job profiles.

[0096] Constraints, as shown in Fig. 59, are generally applied to the system based upon naturally occurring relationships between FieldGroups as they pertain to job profiles in general. For example, it is assumed that the existence of certain FieldGroups presupposes other FieldGroups. A human administrator with expertise in the domain defines these relationships based upon domain knowledge. Because the number of relations between FieldGroups is far fewer than the number of unique survey possibilities, it is efficient for an expert to define these relations, even if there are hundreds of FieldGroups. For instance, a FieldGroup such as "Bar Association" (generally used for lawyers and profiles requiring a legal degree) would only be asked if we knew the user's "Job," or perhaps their "Industry," but it would be pointless to ask that FieldGroup if only the user's "Gender" were known. As such, a FieldGroup such as "Bar Association" is related to "Job" and "Industry," but not to "Gender." Relations such as these are referred to in the system as "FieldGroup Related," and are stored in the system using a relational format. In addition to the levels, these relations act as constraints on the determination of the output of the suggest FieldGroup.

[0097] Given the constraints and weights that are resident in the system, the survey engine selects amongst all the FieldGroups available for the survey the FieldGroup that is the most popular (or highest weighted) FieldGroup among the affinity groups to which the user belongs, constrained by the levels and FieldGroup relations described previously. Finally, the FieldGroup is forwarded to the user for presentation. The user answers it, and the process is repeated for subsequent questions. If, after iteration, a FieldGroup no longer can be suggested that meets the criteria discussed herein, the system marks the users survey as complete and takes a pre-defined next action, such as showing a message to the user and moving on to the reporting aspects of the system as shown in Fig. 26.

[0098] Reference is now made to Figs. 9-10 that show a non-limiting example of the system suggesting popular answers for a FieldGroup in the survey. A mechanism exists in the system, which suggests possible answer choices for a particular

FieldGroup, based upon several factors. These factors are, for example, the FieldGroup being asked, an answer relations document, the user's profile and associated affinity groups, and a set of weights, which store the most popular answers for a particular FieldGroup. Just as there is a relational basis between FieldGroups asked during the survey process, one aspect of the survey engine establishes a relationship between a particular FieldGroup and the suggested answers as shown in Fig. 57. A particular FieldGroup is answer related to another FieldGroup if the answer to the first FieldGroup causes popular answers for the second FieldGroup to be suggested.

[0099]    The 'suggest popular answers' algorithm is constrained by the answer relations, and the answer values are weighted using a table consisting of a list of associations between a value, an affinity group, and the number of profiles who have answered that particular value. Among the constrained answers, the survey engine selects the X most popular (most highly weighted) answers among the list, where X is an integer defined by an administrator as a reasonable number of values to display, from which a user may select.

[0100]    For example if a survey begins with a FieldGroup for Industry, and the FieldGroup is displayed, which asks the user about the Industry in which they work, and they respond "law," a second FieldGroup for Law Firm, based on the process described earlier for suggested questions, is displayed and the user is asked with which law firm they are associated. Based on the Suggest Popular Answers algorithm, the engine can also provide a list of law firms that were previously provided by other users to this FieldGroup, based on those who also identified themselves in the "law" Industry. The affinity group Industry = "law" is, therefore, answer related to the FieldGroup for Law Firm because the suggested answers for the FieldGroup Law Firm are related to the answer provided by the first FieldGroup Industry. Correspondingly, for the FieldGroup Law Firm the system looks up all other FieldGroups that are answer-related to it. For example, the type of job (FieldGroup for Job) may be answer related to the affinity group associated with law firms, as may be the FieldGroup for "total hours billed per year," etc. Therefore, because FieldGroup Job may be answer-related to FieldGroup Industry and FieldGroup Law Firm, the system suggests to the user to select his type of job from a list that includes corporate attorney, litigation attorney, paralegal, and so forth. While, if the user had answered "Computer Software"

for the FieldGroup Industry, then the system would have suggested different possible answers for FieldGroup Job, such as computer programmer, senior software engineer, IT support technician, etc. Alternatively, instead of using the suggested popular answers, the user may enter a new value, and input a job title that is completely new to the system. Based on the user's responses, the system categorizes the users profile and aligns it with profiles that possess similar characteristics.

[0101]   When the survey wizard displays a question, one aspect is to determine if the FieldGroup is an open FieldGroup, supporting open-text answers.  An open FieldGroup is one that allows for free text entries by a user, as opposed to a closed FieldGroup, where a user may only select values from a list or use numeric answers. If the FieldGroup is in an open format, the user may type a open-text answer, or select a suggested popular answer, as described previously. When the user has made their entry, the system determines if the answer was typed as open-text or selected from the list of possible answers.  If the answer is selected from a list of suggested answers, the users choice is saved in a database using the existing Answer ID for that existing answer in the system.  If the answer was free/open-text, an algorithm is invoked to determine if the answer already exists in the database in a similar form to the users free-text entry. Using search technology, possible existing alternatives from the database are suggested that possess similar characteristics as the free-text answer, as shown in Fig. 22. The alternative responses may possess different spelling of key words, grammatical variations, or combinations of other synonymous words. By doing this, the survey wizard identifies the underlying focus of the new answer and ensures that the system is gaining the right perspective from the user. If the user rejects the suggested alternatives, the system requests that the user retype the answer to confirm the response, and possibly enter more information about the answers, so that the system or an administrator may categorize the new answer.  Once the user confirms the response by typing it a second time, the new answer is saved in the database and it becomes one of the possible answers that subsequent users may choose, as shown in Fig. 28.

[0102]   An administrator, or automated program, can also verify that any new answers are qualified values, such as not containing swear-words or other anomalies. An automated program may exist which uses a dictionary, or other software that can verify the validity of the value. If instead of a new open-text answer, the user selects

one of the alternative responses, the selected answer is saved in the database using the existing Answer ID for that answer. Occasionally, the FieldGroup presented to the user is not of an open format, such as a small list of possible responses that do not change (are not open), or a numeric value, such as a salary figure, or a date, or Boolean values such as Yes/No, as further shown in Fig. 25. In numeric or date situations, there are bounds defined as well for answers. If the answer to the FieldGroup fits into the bounds of the FieldGroup definition, the user's response is saved in the database. For instance, if the user is answering a date field, the date may be required to be within a certain range, or if a salary, it must be greater than '0.' If the answer to this closed form of FieldGroup is not within the bounds of the FieldGroup definition, an error message is sent to the user asking for clarification or re-input of the answer. Ultimately, the users profile is saved and cataloged in a database that allows the system to correlate it with other affinity groups and other user profiles to provide a comprehensive compensation report.

[0103] Reference is now made to Fig. 16 where a flow for a profile search is shown. One aspect of the invention concerns comparing an individual profile to all other profiles in the system, and determining which of these profiles are most similar. The Profile Search algorithm returns a list of similar profiles. This similar list of profiles are then summarized, and aggregated into a readable report that offers a complex analysis of a user's compensation and career opportunities. This functionality allows the system to return an in-depth customized report consisting of analysis of similar profiles, in real time. It represents a more accurate picture of a user's compensation than a broad survey could, or one done by human analysts, or based only on job titles, or a standard expert system, etc. The Profile Search algorithm runs, which returns a list of profiles to which the user's profile matches most closely to the search goal. Profiles are retrieved from the database as a set of FieldGroup#,Field#,AnswerValue triplets. The system encapsulates and defines similarity between profiles by use of a scoring mechanism. The scoring mechanism is generic in nature, and can apply to any attribute of a profile that has been defined in the system. It is important to note that different types of similarity can be established by the system. This allows different search goals to be executed. For example, a user may wish to know what projected salary they may expect to make in the marketplace, in their particular occupation, with the all other attributes, such as experience, location and industry, being similar. This is a common

scenario for individuals wishing to switch jobs. The invention supports this type of analysis using a pre-defined search goal document, as shown in Fig. 50, which specifies that the user's profile is to be compared to a set of profiles that have similar jobs, in the same area, industry, etc. In another variation, a user may wish to know what other different jobs for which they may be qualified, given their skills, experience, location, etc. In this case a different search goal is defined for this type of analysis.

[0104] Reference is now made to Figs. 15 and 60, where a profile search scoring system is shown. To encapsulate each type of search goal, a search system is defined according to a relational structure, using a scoring and filtering system. In this scoring system, both exact matches are considered, as well as affinity matches. The affinity matches, although scored lower than exact matches, are critical to the system because they allow for a wide variety of similar attributes to be grouped together, and compared in a more natural way. This solves a critical problem in this area of invention, *i.e.,* previous incarnations of occupational comparison systems require exact matches and, hence, lose much of their value because a large percentage of real world profiles that are similar do not usually have values with exact matches, or the number of true variations is much higher than is usually accounted for in fixed reporting systems. For instance, an IRS auditor may wish to be compared to an accountant, but because they have different job titles, they may never be compared in other systems. In this implementation, affinities (as described elsewhere) are constantly being added and updated through manual and automated processes, which can link these values together and return similar profiles.

[0105] When a profile search is requested, the search goal definition document for the search is retrieved. It specifies rules for how profile attributes are to be matched and compared. The user's profile is then compared to all other profiles in the system, and each matching value from the profiles is assigned a score determined by the field in which it matches. For instance, exact matches on the FieldGroup "job" are assigned a certain relatively high score. An affinity match, *i.e.,* "All People with Accounting Jobs," is also considered because these matches are similar but not exactly the same. Because of this, these matches are given a slightly lower score. Each profile is then scored based on all of its affinity and value matches, and the score is totaled up, by summing up the individual scores. One skilled in the art can understand that many different types of search goals and match criteria can be defined, depending on the

type of FieldGroup. The FieldGroup for "Job" is just one kind. For example, other match scores can be defined for Skills, Specialties, and Certifications, and combined with scores for Job, to create very accurate rankings of profiles returned in the search to meet a given search goal. To be returned for a search (qualify), a profile must meet a certain overall threshold score, which is predefined by an administrator as part of the search goal document. Additionally, the profile must contain matches on specific deterministic fields. The threshold score removes any profiles that do not match on enough attributes to be considered a high quality match. For instance, it is possible that a profile matches on many non-deterministic values, such as gender, geographic location, or experience level, but does not match on something that is critical (or deterministic), such as job type or specialty. An accountant may live next to an attorney, in the same age range, in the same community, who went to the same college, but they should not be considered matches because they are in different jobs and different job affinities. Therefore, in one type of search goal, only certain fields and affinities, such as those related to the job or specialty, are considered deterministic enough to be considered for matches. Again, one skilled in the art will understand that while Job and Skill are considered highly deterministic for basic compensation analysis, the system could be employed for other types of analysis where possibly other FieldGroups such as Location or Gender, might be specified as highly deterministic instead. This allows for the definition and automation of many kinds of reports meeting many different profile search goals.

[0106]    Additionally, some combinations of attribute matches may be considered more valuable than other combinations. For instance, it may be preferable to match a smaller subset of profiles for which a location is closer to the target profile. By employing matchgroups within the profile search goal document, sets of matches can also be ranked. This is useful for instance in ranking profiles higher where geographic proximity is desirable. By creating a set of affinities, described elsewhere, that group together profiles by regions, the profile search system using matchgroups is able to target profiles within local regions first, ranking them higher than profiles that meet other match criteria but are in outlying regions. For instance, the search goal may wish first to find only profiles in the same or surrounding cities. By defining metropolitan region affinity groups, people in the same metropolitan region are found. However, there may not be enough valid data in the system to find profiles in the same

metropolitan region, and therefore, it is also necessary to consider profiles from a larger surrounding region, such as a state, or multi-state area, and choose them if necessary. By including affinities in a ranked fashion in matchgroups, it is possible to return profiles that are closest to a user's by having them rank the highest. One skilled in the art would recognize that closeness could apply to any attribute of a profile in addition to location, such as experience level or age ranges.

[0107]     Reference is now made to Fig. 14 where another aspect of the invention, a set of profiles (usually, but not necessarily, retrieved via the profile search subsystem described elsewhere) is aggregated and their individual attributes are automatically summarized into calculations, such as averages, medians, standard deviations, and counts.   The calculations are aggregated for fast real-time retrieval. Aggregate definitions define dimensions, such as skill, and measures, such as salary, into an aggregate, such as "average salary by skill." Additionally, the system can summarize values into predetermined ranges, such as salary ranges or age ranges. A report chart format defines how charts are displayed to an end user. Many different formats, such as HTML, PDF or JPEG exist.  The report output format may be adjusted to work with any of these formats and these charts may be displayed as bar charts, pie charts, etc.

[0108]     An administrator defines an aggregate definition, as shown in Fig. 41, in an XML format, consisting of measure, a dimension, and a name. Each aggregate is calculated over an affinity group or groups. For example, after a profile search is executed, the list of resultant profiles is combined into an affinity group, such as "People Meeting Search Goal X for User Y," and all available report definitions are executed, resulting in numerous aggregations. Each aggregate definition must contain a measure, a dimension, or both. If no measure is specified, the aggregate is calculated as a count over the entire dimension.  If no dimension is specified, the average, 25th, 50th, 75th median, standard deviation, and standard error, etc, are calculated over the entire population using standard algorithms. Aggregates for any affinity group can be calculated using this method. Using this method allows fast retrieval of aggregate information, and easy definition of new aggregations, which are available to any of thousands of affinity groups defined in the system.

[0109]     Because hundreds of aggregate definitions are available over thousands of affinity groups, it is impossible to present all of this data to a human in a format which they could easily and quickly understand.   The resulting data would consume

thousands of pages. To solve this problem, the invention groups aggregations into charts and groups, and filters and arranges the charts in a layout which is understandable and useful to a human. Only charts that have enough data, and that are determined to be applicable to the user's profile, are shown to the user. This allows for many reports that are only applicable to certain groups to be shown at any time, without having to predefine a report for any particular group. For instance, a group named "Pay Influencers" might contain the following charts: "Average Salary by Job," "Average Salary By Practice Area," "Average Salary By Teaching Rank," "Average Salary By Skill," "Average Salary By Experience," "Average Billing Rate by Bar Association," etc. But not all of the charts in this group are displayed, depending on the applicability to the user's profile. For example, the chart "Average Salary By Teaching Rank" is not displayed if the report is for a lawyer's profile. Another group named "Geographic Outlook" might contain charts such as "Average Salary By City," and "Average Salary By State."

[0110] Charts are grouped into logical sections, as shown in Figs. 29 through 35, that are recognizable to a human, and the system renders the charts in a grouped layout, in an understandable, cohesive, readable presentation. For a particular set of similar profiles, there are many charts that are defined according to a report specification, as shown in Fig. 44. A series of charts are combined into a group, as shown in Fig. 45, which is then parsed and formatted by a program into a user readable format, such as HTML.

[0111] For instance, a lawyer's profile may have matched to many other attorneys, all who have answered a FieldGroup for Bar Association and a FieldGroup for Hourly Billing Rate. The compiled report for the lawyer can automatically show a chart for "Average Billing Rate by Bar Association," whereas a report for a High School Teachers would not show this chart because no matching profiles have either answered the "Bar Association" FieldGroup or the "Hourly Billing Rate" FieldGroup. This is an important innovation over previous inventions, which are restricted to returning generic assessments that group individuals into large, often useless, categorizations while ignoring subtler, yet far more useful categorizations.

[0112] Different embodiments of Report Aggregation and Display exist in the system. "The PayScale Report," as shown in Figs. 29 through 35, is an example of a report aggregation that uses a profile search to define an affinity group for a particular

user, which is then aggregated and displayed according to the logic described herein. The Research Center shown in Fig. 36 is an example of report aggregation and display that uses affinity groups that are defined by affinity definitions, and apply as a general overview of thousands of career related topics. The Profile Alerts shown in Fig. 37 are an example of report aggregation and display that measures similar profile compensation data over time.

[0113] Reference is now made to Fig. 18 where a flow for the rules engine is shown. A rules engine is set up to process profile data automatically to search for common errors, problems, and faulty data. The rules engine is an expert system defined by an expert on the domain. Because the survey system is an open system, which allows free-text user input, as well as freely input numeric data, rules are put in place to monitor that data for validity automatically. This helps to automate the process of data cleaning, and allows an administrator to review large numbers of new profiles more efficiently. For instance, a user may enter data that is obviously bogus, such a having a salary that is too high, or too low for the currency type and country, such as '0.' Also, in some cases the rules engine may also automatically make changes to profile data.

[0114] The rules engine is implemented using a set of database queries, as shown in Fig. 19, and stored procedures that compare the profile to a set of predefined criteria, and that then take some action as a result. Each of these queries is run against each new profile input into the system by users. The results of the rules engine queries, including a list of changes determined by the rules, are stored in a table. The rules engine allows stored procedures to be set up which can check for any type of data error. Another example of a rule is if a user enters too many values for a particular FieldGroup, such as Skill. If a user has answered too many skills, chances are that some of those skills are not valid or the user is just playing with the system, and hence, the profile is invalidated, and not considered for profile searches and reports. A flag on the profile marks a profile as active or inactive. Inactive profiles are not used by any other part of the system for calculations, including comparisons in the survey engine, search goals, or reports. This way, data are cleaned and maintained as statistically valid. Many other flags are also set on a profile, and one skilled in the art can easily recognize that these flags are useful ways to gauge profiles' use in various aspects of the system. Rules in this form can also be used to modify existing profiles for data

entry mistakes. For instance, a user may commonly enter a new open-text value for one FieldGroup that is truly a value for another FieldGroup. To keep data as valid as possible in an open system such as this, the rules engine moves the answer from one FieldGroup to another. One can easily see that a rules engine may be extended to any situation that applies to a significant number of profiles.

[0115]    Figs. 38 through 60 provide XML code and are intended as an exemplary and non-limiting code examples. A person skilled in the art could easily adapt these code examples as may be deemed necessary and remain within the intended scope of this invention.

[0116]    Although the invention is described herein with reference to the preferred embodiment, one skilled in the art will readily appreciate that other applications may be substituted for those set forth herein without departing from the spirit and scope of the present invention. Accordingly, the invention should only be limited by the Claims included below.